AUTOMATA STUDIES

W. R. ASHBY
J. T. CULBERTSON
M. D. DAVIS
S. C. KLEENE
K. DE LEEUW
D. M. MAC KAY
J. MC CARTHY
M. L. MINSKY
E. F. MOORE
C. E. SHANNON
N. SHAPIRO
A. M. UTTLEY
J. VON NEUMANN

Edited by

C. E. Shannon and J. McCarthy

ANNALS OF MATHEMATICS STUDIES

PRINCETON UNIVERSITY PRESS

Annals of Mathematics Studies
Number 34

AUTOMATA STUDIES

W. R. ASHBY
J. T. CULBERTSON
M. D. DAVIS
S. C. KLEENE
K. DE LEEUW
D. M. MAC KAY
J. MC CARTHY
M. L. MINSKY
E. F. MOORE
C. E. SHANNON
N. SHAPIRO
A. M. UTTLEY
J. VON NEUMANN

Edited by
C. E. Shannon and J. McCarthy

Princeton, New Jersey
Princeton University Press

Copyright © 1956, by Princeton University Press
All Rights Reserved

L.C. Card 56-7637
ISBN 0-691-07916-1

Second Printing, 1957
Third Printing, 1960
Fourth Printing, 1965
Fifth Printing, 1972

This research was supported in part by the Office of Naval Research. Reproduction, translation, publication, use and disposal in whole or in part by or for the United States Government is permitted.

Printed in the United States of America

PREFACE

Among the most challenging scientific questions of our time are the corresponding analytic and synthetic problems: How does the brain function? Can we design a machine which will simulate a brain? Speculation on these problems, which can be traced back many centuries, usually reflects in any period the characteristics of machines then in use. Descartes, in DeHomine, sees the lower animals and, in many of his functions, man as automata. Using analogies drawn from water-clocks, fountains and mechanical devices common to the seventeenth century, he imagined that the nerves transmitted signals by tiny mechanical motions. Early in the present century, when the automatic telephone system was introduced, the nervous system was often likened to a vast telephone exchange with automatic switching equipment directing the flow of sensory and motor data. Currently it is fashionable to compare the brain with large scale electronic computing machines.

Recent progress in various related fields leads to an optimistic view toward the eventual and not too remote solution of the analytic and synthetic problems. The fields of neuro-anatomy and neuro-physiology, using powerful tools from electronics in encephalographic studies, have illuminated many features of the brain's operation.

The development of large scale computers has led to a clearer understanding of the theory and design of information processing devices. Programming theory, still in its infancy, is already suggesting the tremendous potential versatility of computers. Electronic computers are also valuable tools in the study of various proposed models of nervous systems. Often such models are much too complex for analytic appraisal, and the only available method of study is by observation of the model in operation.

On the mathematical side, developments in symbolic logic, recursive function theory and Turing machine theory have led to deeper understanding of the nature of what is computable by machines. Neighboring fields of game theory and information theory have developed concepts relating to the nature and coding of information that may prove useful in automata research.

The problem of giving a precise definition to the concept of "thinking" and of deciding whether or not a given machine is capable of thinking has aroused a great deal of heated discussion. One interesting definition has been proposed by A. M. Turing: a machine is termed capable of thinking if it can, under certain prescribed conditions, imitate a human being by answering questions sufficiently well to deceive a human questioner for a reasonable period of time. A definition of this type has the advantages of being operational or, in the psychologists' term, behavioristic. No metaphysical notions of consciousness, ego and the like are involved. While certainly no machines at the present time can even make a start at satisfying this rather strong criterion, Turing has speculated that within a few decades it will be possible to program general purpose computers in such a way as to satisfy this test.

A disadvantage of the Turing definition of thinking is that it is possible, in principle, to design a machine with a complete set of arbitrarily chosen responses to all possible input stimuli (see, in this volume, the Culbertson and the Kleene papers). Such a machine, in a sense, for any given input situation (including past history) merely looks up in a "dictionary" the appropriate response. With a suitable dictionary such a machine would surely satisfy Turing's definition but does not reflect our usual intuitive concept of thinking. This suggests that a more fundamental definition must involve something relating to the manner in which the machine arrives at its responses--something which corresponds to differentiating between a person who solves a problem by thinking it out and one who has previously memorized the answer.

The present volume is a collection of papers which deal with various aspects of automata theory. This theory is of interest to scientists in many different fields and, correspondingly, among the authors are workers who are primarily logicians, mathematicians, physicists, engineers, neurologists and psychologists. The papers include some which are close to pure mathematics; others are essentially directed to the synthesis problem and some relate largely to philosophic questions. There is also a certain amount of overlap, the same problem being handled from somewhat different points of view by different authors.

The papers have been divided into three groups. The first group consists of papers dealing with automata having a finite number of possible internal states. In the usual quantized model of this type, the automaton has a finite number of inputs and outputs and operates in a quantized time scale. Thus, such a device is characterized by two functions of the current state and input, one function giving the next state and the other the next output. Although seemingly trivial at this level of description, many interesting problems arise in the detailed analysis of such machines. Indeed, it should be remembered that essentially all actual physical machines and even the brain itself are, or can be reasonably idealized to be, of this form.

Neurophysiologists have proposed a number of models for the neuron and Kleene, in his paper, investigates the capabilities and limitations of automata constructed from these idealized components. von Neumann, using similar components, allows the possibility of statistically unreliable operation and shows that under certain conditions it is possible, with unreliable components, to construct large and complex automata with as high a reliability as desired. In Culbertson's paper a simple construction is given for an idealized neural network which will react in an arbitrary prescribed manner for an arbitrary lifetime. In all of these papers the notion of universal components plays a significant role. These are components which, roughly speaking, are sufficiently flexible to form devices capable of acting like any machine. Minsky considers the problem of universality of components

and finds conditions which ensure this property. In a paper of a somewhat different type, Moore studies what can be learned about finite state automata by experiments performed on the inputs and outputs of the machine (without direct observation of its interior).

The second group of papers deals with the theory of Turing machines and related questions, that is to say, with automata having an unlimited number of possible states. The original Turing machine (since then, recast in many different forms) may be described as follows. Let there be given a tape of infinite length which is divided into squares and a finite list of symbols which may be written on these squares. There is an additional mechanism, the head, which may read the symbol on a square, replace it by another or the same symbol and move to the adjoining square to the left or right. This is accomplished as follows: At any given time the head is in one of a finite number of internal states. When it reads a square it prints a new symbol, goes into a new internal state and moves to the right or left depending on the original internal state and the symbol read. Thus a Turing machine is described by a finite list of quintuplets such as 3, 4, 3, 6, R which means: If the machine is in the third internal state and reads the fourth symbol it prints the third symbol, goes into the sixth internal state and moves to the right on the tape. There is a fixed initial internal state and the machine is supposed to start on a blank tape. One of the symbols represents a blank square and there may be given a state in which the machine stops.

Turing gave a convincing argument to the effect that any precisely defined computation procedure could be carried out by a machine of the type described above. He also showed that the Turing machines can be enumerated and that a universal machine could be made which, when it read the number of any Turing machine, would carry out the computation that that machine would have carried out were it put on a blank tape. His final result was to show that there did not exist a Turing machine which when confronted with the number of another machine would decide whether that machine would ever stop.

Any of the present automatic electronic computers is equivalent to a universal Turing machine if it is given, for example, a means of asking for more punched cards and for the return of cards it has already punched. In Shannon's paper it is shown that a universal Turing machine can be constructed with only two internal states, or alternatively, with only two tape symbols. Davis gives a general definition of a universal Turing machine and establishes some results to make this definition appear reasonable. McCarthy discusses the problem of calculating the inverse of the function generated by a Turing machine, after some argument to the effect that many intellectual problems can be formulated in this way. Finally, De Leeuw, Moore, Shannon and Shapiro investigate whether machines with random elements can compute anything uncomputable by ordinary Turing machines.

The third section of the book contains papers relating more directly to the synthesis of automata which will simulate in some sense the operation of a living organism. Ashby discusses the problem of designing an intelligence amplifier, a device which can solve problems beyond the capacities of its designer. MacKay, dealing with the same general problem, suggests means for an automaton to symbolize new concepts and generate new hypotheses. Uttley studies from a still different point of view the problem of the abstraction of temporal and spatial patterns by a machine, that is, the general problem of concept formation.

It gives us pleasure to express our gratitude to all those who have contributed to the preparation of this volume. The work was supported in part by the Princeton Logistics Project sponsored by the Office of Naval Research. Professor A. W. Tucker, directing this project, has been most helpful. H. S. Bailey, Jr. and the staff of the Princeton University Press, particularly Mrs. Dorothy Stine and Mrs. Jean Muiznieks have been efficient and cooperative. Thanks are also due Dr. Julia Robinson for help with the reviewing and Mrs. E. Powanda for secretarial services.

John McCarthy
Claude Shannon

CONTENTS

Preface v

FINITE AUTOMATA

Representation of Events in Nerve Nets and Finite Automata
By S. C. Kleene 3

Probabilistic Logics and the Synthesis of Reliable Organisms From Unreliable Components
By J. von Neumann 43

Some Uneconomical Robots
By James T. Culbertson 99

Some Universal Elements for Finite Automata
By M. L. Minsky 117

Gedanken-Experiments on Sequential Machines
By Edward F. Moore 129

TURING MACHINES

A Universal Turing Machine with Two Internal States
By Claude E. Shannon 157

A Note on Universal Turing Machines
By M. D. Davis 167

The Inversion of Functions Defined by Turing Machines
By John McCarthy 177

Computability by Probabilistic Machines
By K. de Leeuw, E. F. Moore, C. E. Shannon, and N. Shapiro 183

SYNTHESIS OF AUTOMATA

Design for an Intelligence-Amplifier
By W. Ross Ashby 215

The Epistemological Problem for Automata
By D. M. MacKay 235

Conditional Probability Machines and Conditioned Reflexes
By Albert M. Uttley 253

Temporal and Spatial Patterns in a Conditional Probability Machine
By Albert M. Uttley 277

FINITE AUTOMATA

REPRESENTATION OF EVENTS IN NERVE NETS AND FINITE AUTOMATA[1]

S. C. Kleene

INTRODUCTION

1. Stimuli and Response

An organism or an automaton receives stimuli via its sensory receptor organs, and performs actions via its effector organs. To say that certain actions are a response to certain stimuli means, in the simplest case, that the actions are performed when and only when those stimuli occur.

In the general case both the stimuli and the actions may be very complicated.

In order to simplify the analysis, we may begin by leaving out of account the complexities of the response. We reason that any sort of stimulation, or briefly any event, which affects action in the sense that different actions ensue according as the event occurs or not, under some set of other circumstances held fixed, must have a representation in the state of the organism or automaton, after the event has occurred and prior to the ensuing action.

So we ask what kind of events are capable of being represented in the state of an automaton.

For explaining actions as responses to stimuli it would remain to study the manner in which the representations of events (a kind of internal response) lead to the overt responses.

Our principal result will be to show (in Sections 7 and 9) that all and only the events of a certain class called "regular events" are representable.

---

[1] The material in this article is drawn from Project RAND Research Memorandum RM-704 (15 December 1951, 101 pages) under the same title and by the author. It is used now by permission of the RAND Corporation. The author's original work on the problem was supported by the RAND Corporation during the summer of 1951.

2. Nerve Nets and Behavior

McCulloch and Pitts [1943] in their fundamental paper on the logical analysis of nervous activity formulated certain assumptions which we shall recapitulate in Section 3.

In showing that each regular event is representable in the state of a finite automaton, the automaton we use is a McCulloch-Pitts nerve net. Thus their neurons are one example of a kind of "universal elements" for finite automata.

The McCulloch-Pitts assumptions were put forward as an abstraction from neuro-physiological data. We shall not be concerned with the question of how exactly the assumptions fit. They seem to fit roughly up to a point, though one of McCulloch's and Pitts' results is that certain alternative assumptions can explain the same kind of behavior. With increasing refinement in the neuro-physiological data the emphasis is no doubt on respects in which the assumptions do not fit.

Our theoretical objective is not dependent on the assumptions fitting exactly. It is a familiar strategem of science, when faced with a body of data too complex to be mastered as a whole, to select some limited domain of experiences, some simple situations, and to undertake to construct a model to fit these at least approximately.

Having set up such a model, the next step is to seek a thorough understanding of the model itself. It is not to be expected that all features of the model will be equally pertinent to the reality from which the model was abstracted. But after understanding the model, one is in a better position to see how to modify or adapt it to fit the limited data better or to fit a wider body of data and when to seek some fundamentally different kind of explanation.

McCulloch and Pitts in their original paper give a theory for nerve nets without circles [Part II of their paper] and a theory for arbitrary nerve nets [Part III]. The present article is partly an exposition of their results; but we found the part of their paper dealing with arbitrary nerve nets obscure, so we have proceeded independently there.

Although we are concerned with the model itself rather than its application, a few remarks on the latter may prevent misunderstanding.

To take one example, as consideration of the model shows, memory can be explained on the basis of reverberating cycles of nerve impulses. This seems a plausible explanation for short-term memories. For long-term memories, it is implausible on the ground of fatigue, also on the ground that calculations on the amount of material stored in the memory would call for too many neurons [McCulloch, 1949], and also on the basis of direct experimental evidence that temporary suppression of nervous activity does not cut off memory [Gerard, 1953].

The McCulloch-Pitts assumptions give a nerve net the character of a digital automaton, as contrasted to an analog mechanism in the sense familiar in connection with computing machines. Some physiological processes of control seem to be analog. Just as in mathematics continuous processes can be approximated by discrete ones, analog mechanisms can be approximated in their effect by digital ones. Nevertheless, the analog or partly analog controls may for some purposes be the simplest and most economical.

An assumption of the present mathematical theory is that there are no errors in the functioning of neurons. Of course this is unrealistic both for living neurons and for the corresponding units of a mechanical automaton. It is the natural procedure, however, to begin with a theory of what happens assuming no malfunctioning. Indeed in our theory we may represent the occurrence of an event by the firing of a single neuron. Biologically it is implausible that important information should be represented in an organism in this way. But by suitable duplication and interlacing of circuits, one could then expect to secure the same results with small probability of failure in nets constructed of fallible neurons.

Finally, we repeat that we are investigating McCulloch-Pitts nerve nets only partly for their own sake as providing a simplified model of nervous activity, but also as an illustration of the general theory of automata, including robots, computing machines and the like. What a finite automaton can and cannot do is thought to be of some mathematical interest intrinsically, and may also contribute to better understanding of problems which arise on the practical level.

PART I: NERVE NETS

3. McCulloch-Pitts Nerve Nets

Under the assumptions of McCulloch and Pitts [1943], a nerve cell or *neuron* consists of a body or *soma*, whence nerve fibers (*axons*) lead to one or more *endbulbs*.

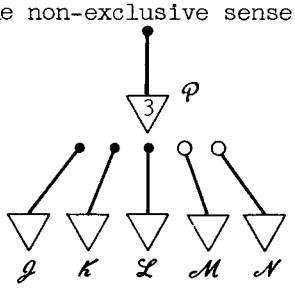
A *nerve net* is an arrangement of a finite number of neurons in which each endbulb of any neuron is adjacent to (*impinges* on) the soma of not more than one neuron (the same or another); the separating gap is a *synapse*. Each endbulb is either *excitatory* or *inhibitory* (not both).

We call the neurons (zero or more) on which no endbulbs impinge *input neurons*; the others, *inner neurons*.

At equally separated moments of time (which we take as the integers on a time scale, the same for all neurons in a given net), each neuron of the net is either *firing* or *not firing* (being *quiet*). For an input neuron, the firing or non-firing at any moment t is determined by conditions outside the net. One can suppose each is impinged on by a sensory receptor organ, which under suitable conditions in the environment causes the neuron to fire at time t. For an inner neuron, the condition for firing at time t

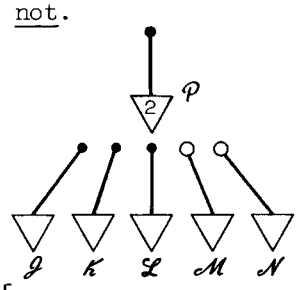
is that at least a certain number h (the threshold of that neuron) of the excitatory endbulbs and none of the inhibitory endbulbs impinging on it belong to neurons which fired at time t - 1.

For illustration, consider the nerve net shown in Figure 1, with input neurons $\mathcal{J}$, $\mathcal{K}$, $\mathcal{L}$, $\mathcal{M}$ and $\mathcal{N}$, and inner neuron $\mathcal{P}$. Excitatory endbulbs are shown as dots, and inhibitory as circles. The threshold of $\mathcal{P}$ is 3, as shown by the number in the triangle representing its soma. The formula written below the net expresses in logical symbolism that neuron $\mathcal{P}$ fires at time t, if and only if all of $\mathcal{J}$, $\mathcal{K}$ and $\mathcal{L}$ and none of $\mathcal{M}$ and $\mathcal{N}$ fired at time t - 1. We are writing "P(t)" to say that neuron $\mathcal{P}$ fires at time t, "J(t - 1)" to say that $\mathcal{J}$ fired at t - 1, etc. The symbol "≡" means if and only if (or is equivalent to), "&" means and, "∨" means or (in the non-exclusive sense), and "¯" means not.



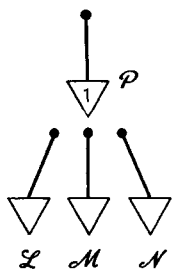
$$P(t) \equiv J(t-1) \,\&\, K(t-1) \,\&\, L(t-1) \,\&\, \overline{M(t-1)} \,\&\, \overline{N(t-1)}.$$

FIGURE 1: Conjunctive Net



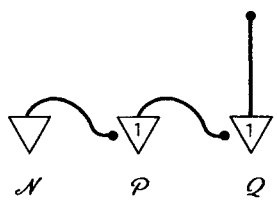
$$P(t) \equiv \Big[[J(t-1) \,\&\, K(t-1)] \vee [J(t-1) \,\&\, L(t-1)] \vee [K(t-1) \,\&\, L(t-1)]\Big] \,\&\, \overline{M(t-1)} \,\&\, \overline{N(t-1)}.$$

FIGURE 2



$$P(t) \equiv L(t-1) \vee M(t-1) \vee N(t-1).$$

FIGURE 3: Disjunctive Net



$$P(t) \equiv N(t-1),$$

$$Q(t) \equiv N(t-2).$$

FIGURE 4: Delay Net

4. The Input to a Nerve Net

Consider a nerve net with k input neurons $\mathcal{N}_1, \ldots, \mathcal{N}_k$. We assume $k \geq 1$, until 6.3. The input (or experience) over all past time up to the present moment inclusive can be described by a table with k columns

corresponding to the input neurons, and with rows corresponding to the moments counting backward from the present moment p. The positions are filled with 0's and 1's, where 0 is to stand for quiescence, and 1 for firing, of the neuron in question at the moment in question.

For example with k = 2 the table might be like Figure 5. The 1 in the first row and first column means that $\mathcal{N}_1$ fires at time p, the 0 in the third row and first column that $\mathcal{N}_1$ did not fire at time p - 2, etc. If this table is extended down infinitely, we have a representation of the input, thought of as extending over all past time, which for the time being we treat as infinite. (In Section 6 we shall reconsider the matter.)

t	$\mathcal{N}_1$	$\mathcal{N}_2$
p	1	0
p - 1	1	1
p - 2	0	1
. . .		

FIGURE 5

By an event we shall mean any property of the input. In other words, any subclass of the class of all the possible tables describing the input over all past time (and ending with the present p inclusive, except when otherwise stated) constitutes an event, which occurs when the table describing the actual input belongs to this subclass.

Examples of events with two input neurons $\mathcal{N}_1$ and $\mathcal{N}_2$ are:

(1) $\mathcal{N}_1$ fires at time p.
(2) $\mathcal{N}_2$ does not fire at time p, and $\mathcal{N}_1$ fired at time p - 1.
(3) One of $\mathcal{N}_1$ and $\mathcal{N}_2$ fires at time p.
(4) $\mathcal{N}_1$ and $\mathcal{N}_2$ both fire at time p.
(5) $\mathcal{N}_2$ fired at some time.
(6) $\mathcal{N}_2$ fired at every time except p.

Of these, the input described by the table of Figure 5 constitutes an occurrence of events (1), (2), (3) and (5), but not of (4), while we need to know the rest of the table to know whether it constitutes an occurrence of (6).

5. Definite Events

5.1 **"DEFINITE EVENTS" DEFINED.** We shall discuss first events which refer to a fixed period of time, consisting of some ℓ (≥ 1) consecutive moments p - ℓ + 1, ..., p ending with the present. We call such events definite of length (or duration) ℓ. Of the preceding examples, (1) – (4) are definite, but not (5) and (6).

Then in a table such as Figure 5 we need consider only the uppermost ℓ rows; e.g., that table for $\ell = 3$ then describes an event also described by the formula $N_1(p)\ \&\ \overline{N_2(p)}\ \&\ N_1(p-1)\ \&\ N_2(p-1)\ \&\ \overline{N_1(p-2)}\ \&\ \&\ N_2(p-2)$.

There are exactly $k\ell$ entries in a table describing the input on k neurons for the ℓ moments p - ℓ + 1, ..., p. Therefore there are exactly $2^{k\ell}$ possible such tables. Therefore there are exactly $2^{2^{k\ell}}$ definite

events on k input neurons of length ℓ, since any particular one is determined by saying which of the inputs described by the $2^{k\ell}$ $k \times \ell$ tables would constitute an occurrence of the event.

We call a definite event positive, if it occurs only when at least one input neuron fires during the period to which the event refers. There are exactly $2^{2^{k\ell}-1}$ positive definite events on k input neurons of length ℓ, since that input described by the table of all 0's is excluded as an occurrence.

5.2 REPRESENTABILITY OF DEFINITE EVENTS: AN ILLUSTRATION. Consider the definite event which occurs when the pattern of firings fits either the table of Figure 5 (stopped at three rows) or that of Figure 6.

t	$\mathcal{N}_1$	$\mathcal{N}_2$	
p	1	0	$N_1(p)$ & $\overline{N_2(p)}$
p - 1	1	0	& $N_1(p-1)$ & $\overline{N_2(p-1)}$
p - 2	1	0	& $N_1(p-2)$ & $\overline{N_2(p-2)}$.

FIGURE 6

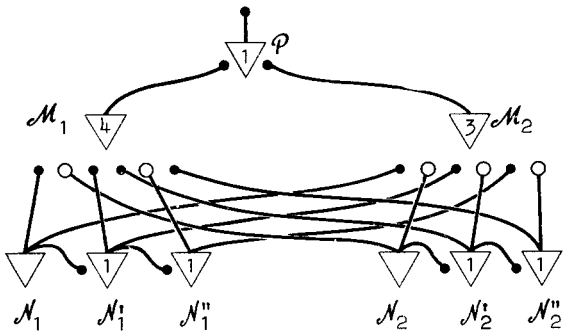
That is, exactly these two (out of the $2^{2 \cdot 3} = 64$) 2×3 tables are to constitute an occurrence of the event. The event is described by the right member of the equivalence in Figure 7, which is obtained by combining disjunctively the conjunctions describing the respective tables separately. In the nerve net of Figure 7, the neuron $\mathcal{P}$ fires at time $p + 2$, if and only if the event occurs ending at time p; or briefly, the net represents the event by the firing of $\mathcal{P}$ with lag 2. The neurons $\mathcal{N}_1$, $\mathcal{N}_1'$, $\mathcal{N}_1''$ with just the axons connecting them are a "delay net" (cf. Figure 4). The synapse at $\mathcal{M}_1$ with the seven neurons involved is a "conjunctive net" (cf. Figure 1). That at $\mathcal{P}$ is a "disjunctive net" (cf. Figure 3).

The method of this illustration applies to every positive definite event which occurs for some one or more tables.

There remains the case of the event which never occurs. This is represented by the firing of $\mathcal{P}$ say with lag 2 in the net of Figure 8. The neuron $\mathcal{M}_2$ is inserted to show that we can have the net connected (in the obvious sense); otherwise $\mathcal{M}_2$ could be omitted. $\mathcal{M}_1$ could be the $\mathcal{P}$.

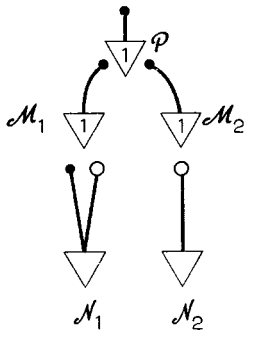
We have thus already proved that any positive definite event is representable by firing a neuron with lag 2. However, we shall give a more flexible treatment, listing this result as part of Theorem 1 Corollary 1.

5.3 REPRESENTABILITY OF DEFINITE EVENTS: GENERAL THEORY. We consider logical expressions constructed using & and $\vee$ from the expressions symbolizing



$$P(p+2) \equiv \Big[N_1(p) \,\&\, \overline{N_2(p)} \,\&\, N_1(p-1) \,\&\, N_2(p-1) \,\&\, \overline{N_1(p-2)} \,\&\, N_2(p-2)\Big] \vee \Big[N_1(p) \,\&\, \overline{N_2(p)} \,\&\, N_1(p-1) \,\&\, \overline{N_2(p-1)} \,\&\, N_1(p-2) \,\&\, \overline{N_2(p-2)}\Big].$$

FIGURE 7



$\mathcal{P}$ does not fire at time $p+2$, or in symbols $P(p+2) \equiv N_1(p) \,\&\, \overline{N_1(p)}$.

FIGURE 8

the firing or non-firing of one of the k input neurons $\mathcal{N}_1, \ldots, \mathcal{N}_k$ at one of the ℓ moments $p - \ell + 1, \ldots, p$. Such an expression we call a $k\ell$ - formula, of depth equal to the greatest number of successive times & or $\vee$ is used in its construction. Here we allow any bracketing in conjunctions and disjunctions of more than two members; so e.g., $[[N_1(p) \,\&\, \overline{N_2(p)}] \,\&\, N_2(p-1] \vee \overline{N_1(p)}$ as written is of depth 3, but it can be rewritten as $[N_1(p) \,\&\, \overline{N_2(p)} \,\&\, N_2(p-1)] \vee \overline{N_1(p)}$ with depth 2.

These definitions can be given by mathematical induction on s, thus:

(1) For each i and j $(i = 1, \ldots, k;\ j = 1, \ldots, \ell)$, $N_i(p-j+1)$ and $\overline{N_i(p-j+1)}$ are $k\ell$ - formulas of depth 0.

(2) For $s > 0$, if $G_1, \ldots, G_n$ $(n \geq 2)$ are $k\ell$ - formulas of depths $< s$, at least one of them being of depth $s-1$, then $G_1 \,\&\, \ldots \,\&\, G_n$ and $G_1 \vee \ldots \vee G_n$ are $k\ell$ - formulas of depth s. (Here each G_e not of depth 0 is to be enclosed in brackets when written out.)

Since the truth or falsity of a $k\ell$ - formula F is determined logically from only the truth or falsity of the $N_i(p-j+1)$ which enter into it as prime components, each F expresses a definite event E on k input neurons of length ℓ.

We lose nothing essential by applying the negation symbol $\overline{}$ only directly to the prime components. For by repeated use of the logical identities $\overline{\overline{G}} \equiv G$, $\overline{G_1 \,\&\, \ldots \,\&\, G_n} \equiv \overline{G_1} \vee \ldots \vee \overline{G_n}$ and $\overline{G_1 \vee \ldots \vee G_n} \equiv \overline{G_1} \,\&\, \ldots \,\&\, \overline{G_n}$, negation symbols used otherwise could be moved inward to the prime components without changing the depth. The only other operations commonly employed in the

two-valued propositional calculus, namely $\rightarrow$ (implies) and $\equiv$, can be expressed in terms of $\overline{}$, $\&$ and $\vee$ thus: $G \rightarrow H \equiv \bar{G} \vee H$, $(G \equiv H) \equiv (G \rightarrow H) \,\&\, (H \rightarrow G)$.

A circle (of length c) in a nerve net is a set of distinct neurons $\mathcal{N}_1, \ldots, \mathcal{N}_c$ $(c \geq 1)$ such that $\mathcal{N}_i$ has an endbulb on $\mathcal{N}_{i+1}$ for each i $(i = 1, \ldots, c - 1)$ and $\mathcal{N}_c$ has an endbulb on $\mathcal{N}_1$. The nets so far considered are without circles, including the conjunctive, disjunctive and delay nets (Figures 1, 3 and 4), and certain nets composed thence.

> THEOREM 1. Let F be any $k\ell$ - formula of depth s, and let E be the definite event on k input neurons of length ℓ which F expresses. There is a nerve net of structure corresponding to F (and therefore without circles) which represents E by firing or by not firing, according as E is positive or non-positive, a certain neuron $\mathcal{P}$ (inner if $s > 0$) at time $p + s$.

By saying that the net is of "structure corresponding to F", we mean that it is composed out of conjunctive and disjunctive nets (together with delay nets) corresponding to the operations used in constructing F, as will be indicated in the proof.

PROOF, by induction on s. Under our definition of $k\ell$ - formula not all of the symbols N_i $(i = 1, \ldots, k)$ need occur in F. In showing by induction how to construct the net to correspond to the logical structure of F, we incorporate only the neurons $\mathcal{N}_i$ for which N_i occurs in F. The others can be considered as floating around, unless one wishes them connected to the rest of the net, in which case if $s > 1$ they can be, e.g., as illustrated for $\mathcal{N}_2$ in Figure 8.

BASIS: $s = 0$. Then F is $N_i(p - j + 1)$ or $\overline{N_i(p - j + 1)}$ for some i and j. Then $\mathcal{N}_i$ is also the $\mathcal{P}$, if $j = 1$; and otherwise the $\mathcal{P}$ is a neuron coming from $\mathcal{N}_i$ by a suitable delay net (Figure 4).

INDUCTION STEP: $s > 0$. Then F is $G_1 \,\&\, \ldots \,\&\, G_n$ or $G_1 \vee \ldots \vee G_n$. For $e = 1, \ldots, n$, let M_e be G_e or $\overline{G_e}$ according as (the event described by) G_e is positive or non-positive, as will be known from the case which applied to G_e. Then G_e (of depth $s_e < s$) is equivalent to M_e or $\overline{M}_e$, respectively, and by the hypothesis of the induction, there is a nerve net with a neuron $\mathcal{G}_e$ the firing or non-firing, respectively, of which at time $p + s_e$ represents G_e. Thence we obtain a neuron $\mathcal{M}_e$ the firing of which at time $p + s - 1$ represents M_e; this $\mathcal{M}_e$ is $\mathcal{G}_e$ itself if $s_e = s - 1$, and otherwise a neuron coming from $\mathcal{G}_e$ by a suitable delay

net. Now we have four cases, according to how F is composed out of $M_1, \ldots, M_n$, via the construction of F from $G_1, \ldots, G_n$ and the equivalence of each G_e to one of M_e and $\overline{M}_e$.

CASE 1: A conjunction containing at least one unnegated factor, e.g., $M_1 \,\&\, M_2 \,\&\, M_3 \,\&\, \overline{M}_4 \,\&\, \overline{M}_5$. The event is then positive; so we wish to represent it by the firing of a neuron $\mathcal{P}$ at time $p + s$. A conjunctive net (cf. Figure 1) gives us this neuron.

CASE 2: A conjunction containing only negated factors, e.g., $\overline{M}_1 \,\&\, \overline{M}_2 \,\&\, \overline{M}_3$. The event is then non-positive. But its negation $\overline{\overline{M}_1 \,\&\, \overline{M}_2 \,\&\, \overline{M}_3}$ is positive. The latter is equivalent to $M_1 \vee M_2 \vee M_3$. A disjunctive net (cf. Figure 3) represents the latter by firing a neuron $\mathcal{P}$ at $p + s$; the net then represents the original event by the non-firing of $\mathcal{P}$ at $p + s$, which is how we wished it to be represented.

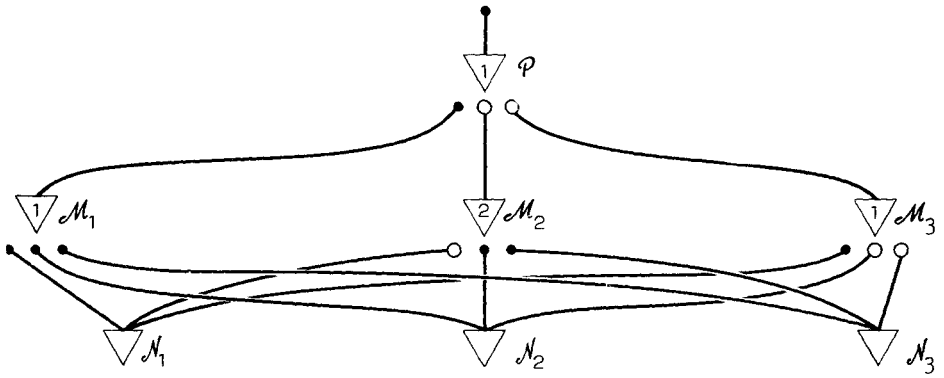
CASE 3: A disjunction containing at least one negated term, e.g., $\overline{M}_1 \vee \overline{M}_2 \vee \overline{M}_3 \vee M_4 \vee M_5$. The event is non-positive. But its negation is positive and equivalent to $M_1 \,\&\, M_2 \,\&\, M_3 \,\&\, \overline{M}_4 \,\&\, \overline{M}_5$. A conjunctive net represents the latter by firing a neuron $\mathcal{P}$ at $p + s$; then the original event is represented by the non-firing of $\mathcal{P}$ at $p + s$.

CASE 4: A disjunction containing only unnegated terms, e.g., $M_1 \vee M_2 \vee M_3$. The event is positive. A disjunctive net represents it as desired by firing a neuron $\mathcal{P}$ at $p + s$.

EXAMPLES. The net of Figure 7 is what the present method gives for the formula there. Another illustration is in Figure 9. Treating the three formulas $N_1(p) \vee N_2(p) \vee N_3(p)$, $\mathcal{N}_1(p) \vee \overline{\mathcal{N}_2(p)} \vee \overline{\mathcal{N}_3(p)}$ and $\overline{\mathcal{N}_1(p)} \vee \mathcal{N}_2(p) \vee \mathcal{N}_3(p)$ gives us respective neurons $\mathcal{M}_1, \mathcal{M}_2$ and $\mathcal{M}_3$ which represent the events expressed, the first which is positive by firing (Case 4), the second and third which are non-positive by non-firing (Case 3), at time $p + 1$. Then $\mathcal{P}$ is obtained to represent the entire event which is positive by firing at time $p + 2$ (Case 1).

> COROLLARY 1. To each positive (non-positive) definite event, there is a nerve net without circles which represents the event by firing (not firing) a certain inner neuron at time $p + 2$.

The result was stated (for positive events and without the remark on the lag) by McCulloch and Pitts [1943].



$$P(p+2) \equiv \left[N_1(p) \vee N_2(p) \vee N_3(p)\right] \;\&\; \left[N_1(p) \vee \overline{N_2(p)} \vee \overline{N_3(p)}\right] \;\&\; \left[\overline{N_1(p)} \vee N_2(p) \vee N_3(p)\right].$$

FIGURE 9

PROOF. To infer this from the theorem, we need merely observe that the method of 5.2 gives a $k\ell$ - formula of depth ≤ 2 to every definite event on k input neurons of length ℓ. (If the depth is < 2, a delay net may be used to increase the lag to 2.)

DISCUSSION. Readers familiar with symbolic logic will recognize the $k\ell$ - formula so obtained as a principal disjunctive normal form of Hilbert and Ackermann [1928]; it is "principal" because in each of its terms every one of $N_1(p)$, ..., $N_k(p - \ell + 1)$ occurs negated or unnegated (with an exception in the case of Figure 8).

The formula of Figure 9 is a principal conjunctive normal form. If the *p.d.n.f.* has $n < 2^{k\ell}$ terms, the *p.c.n.f.* has $2^{k\ell} - n$ factors. (The *p.c.n.f.* is obtained by evaluating the negation of the *p.d.n.f.* of the negation of the event.)

The use of the *p.d.n.f.* simplifies the proof of representability (cf. 5.2), and gives the fact that the lag can be held to 2, but the net constructed may be unnecessarily complicated. The event may admit of being described more simply by a disjunctive or conjunctive normal form not principal (which still enables the lag to be held to 2). For example (with $k = 2$, $\ell = 3$), $\left[\overline{N_2(p)} \;\&\; N_2(p-2)\right] \vee N_2(p-1)$ is a *d.n.f.*, the *p.d.n.f.* for which would have 40 terms (the *p.c.n.f.* 24 factors). There may be simpler equivalents not disjunctive or conjunctive normal forms.

Since the theorem gives a representing net of corresponding structure to the formula, the problem of finding as simple nets as possible to represent definite events is correlated to the problem of finding simplest equivalents of an expression in the propositional calculus, which has recently been treated by Quine [1952].

In special cases the net may be constructed more simply than corresponding to the formula; e.g., in Figure 2 taking $p = t - 1$ the net represents the event with lag 1, although the formula is of depth 3, and no equivalent formula of depth 1 exists.

Reduction of the lag below 2 is not possible in general. For example (with $k = 3$, $\ell = 1$) the event $N_1(p)\ \&\ (\overline{N_2(p)} \vee \overline{N_3(p)})$ is not representable with lag 1. For it is easily seen that no net consisting of a neuron $\mathcal{P}$ impinged on only by endbulbs belonging directly to $\mathcal{N}_1$, $\mathcal{N}_2$ and $\mathcal{N}_3$ can represent this event.

To hold the lag to 2, we may be obliged to have very large numbers of endbulbs originating from or impinging on a given soma.

> COROLLARY 2. To each positive (non-positive) definite event, there is a number s and a nerve net without circles, composed of neurons each having, and impinged upon by, at most two endbulbs and of threshold at most 2, which represents the event by the firing (non-firing) of a certain neuron at time $p + s$.

PROOF. By using & and $\vee$ only as binary operations in the $k\ell$ - formula, no neuron in the net construction for the theorem will be impinged upon by more than two endbulbs. Each inner neuron outside of the delay nets has only one endbulb. Each input neuron and delay net can if necessary be replaced (increasing the lag) by a tree of neurons each with at most two endbulbs.

We have been considering representation of an event ending with time p by the firing or by the non-firing of a certain neuron at a certain time $p + s$ $(s \geq 0)$. More generally we can consider representation by a property of the state of the net (i.e., the firing or non-firing of each of its neurons) at $p + s$; i.e., the state of the net is to have or not have this property at time $p + s$, according as the event did or did not occur ending with time p. In the following lemma and corollary, it is not being assumed that the event is definite or the net without circles.

> LEMMA 1. An event which is representable in a nerve net by a property of the state at time $p + s$ for a given $s > 0$ is representable by a property of the state of the same net at time p.

PROOF. What happens at times $\leq p$ can only affect the state of the net at time $p + s$ via the state of the entire net, including both the k input neurons and, say, m inner neurons, at time p.

COROLLARY 3. An event which is representable in a nerve net by a property of the state at time $p + s$ for a given $s \geq 0$ is representable by the firing or the non-firing (according to the nature of the property) of a certain inner neuron in a suitable net at time $p + 2$.

PROOF. We can treat the $k + m$ neurons as though all of them were input neurons, for the purpose of applying Corollary 1. By Lemma 1, the property in question is equivalent to a property of the $k + m$ neurons at time p. The latter constitutes a definite event of length 1 on the $k + m$ neurons.

5.4 NERVE NETS WITHOUT CIRCLES.

THEOREM 2. Given any nerve net without circles and given any inner neuron $\mathcal{N}$ in that net, the firing (non-firing) of that neuron at time $p + 1$ is equivalent to the occurrence of a positive (non-positive) definite event.

This theorem was stated (for positive events) by McCulloch and Pitts [1943].

PROOF. Whether or not $\mathcal{N}$ fires at $p + 1$ is completely determined by the state (firing or non-firing) at p of those neurons $\mathcal{N}'_1, \ldots, \mathcal{N}'_r$ having endbulbs impinging on $\mathcal{N}$. Consider those of $\mathcal{N}'_1, \ldots, \mathcal{N}'_r$ which are inner neurons, and repeat the argument. Since there are no circles, any chain of neurons beginning with $\mathcal{N}$ and each impinged on by an endbulb of the next must terminate (with an input neuron). Let $\ell + 1$ be the greatest of the lengths of these chains; since $\mathcal{N}$ is inner, $\ell \geq 1$. After ℓ steps, no inner neurons remain to be considered. Thus whether or not $\mathcal{N}$ fires at time $p + 1$ is completely determined by the state of certain input neurons at certain times between $p - \ell + 1$ and p inclusive; i.e., $\mathcal{N}$'s firing at $p + 1$ is equivalent to a definite event of length ℓ. This event is positive, as firing can only be propagated but not originated under the law for an inner neuron's firing.

REMARK. Any definite event is expressible by a logical formula, e.g., by a principal disjunctive normal form as in 5.2. So a priori there is a formula to express the event of the theorem. By utilizing the condition for firing at each synapse, which can be formulated in logical symbols depending on the threshold and the numbers and kinds of the endbulbs (cf. Figures 1 – 3 for several examples), one can build up a formula directly in ℓ steps, as McCulloch and Pitts indicate.

COROLLARY 1. Any event which is representable in a nerve net without circles by the firing (non-firing) of a given inner neuron $\mathcal{N}$ at time $p + s$ for a given $s \geq 1$ is positive (non-positive) definite.

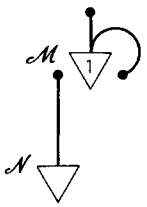
PROOF. By the theorem, the firing of $\mathcal{N}$ at time $p + s$ is equivalent to the occurrence of a positive definite event ending with time $p + s - 1$. But by the hypothesis that $\mathcal{N}$'s firing at time $p + s$ represents an event, i.e., one ending with time p (cf. Section 4), the input over the moments $p + 1, \ldots, p + s - 1$ has no effect on whether $\mathcal{N}$ fires at time $p + s$.

COROLLARY 2. Any event which is representable in a nerve net without circles by a property of the state at time $p + s$ for a given $s \geq 0$ is definite.

PROOF. By Corollary 1, with Theorem 1 Corollary 3 (which does not introduce circles).

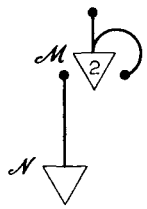
6. Indefinite Events: Preliminaries

6.1 EXAMPLES. Let "(Et)" mean there exists a t ... such that, "(t)" mean for all t, and "$\rightarrow$" mean implies. The net in Figure 10 has a circle of length 1. If at some time $t \leq p$ the input neuron $\mathcal{N}$ fires, then $\mathcal{M}$ will fire at every subsequent moment, in particular at $p + 1$ as the formula expresses. But the firing of $\mathcal{M}$ at time $p + 1$ does not represent the indefinite event $(Et)_{t \leq p} N(t)$ (i.e., we do not have $M(p + 1) \equiv (Et)_{t \leq p} N(t)$), if past time is infinite, because the firing of $\mathcal{M}$ at time $p + 1$ can also be explained by $\mathcal{M}$ having fired at every past moment, without $\mathcal{N}$ having ever fired. Similar examples are given in Figures 11 and 12.



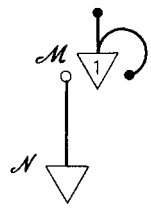
$(Et)_{t \leq p} N(t) \rightarrow M(p + 1)$.

FIGURE 10



$M(p + 1) \rightarrow (t)_{t \leq p} N(t)$.

FIGURE 11



$(Et)_{t \leq p} N(t) \rightarrow \overline{M(p + 1)}$.

FIGURE 12

This difficulty is not escapable by using other nets to represent the events, or in other examples of indefinite events, but constitutes the general rule, by Theorem 6 in Section 10 below with Lemma 1.

Of course any living organism or constructed robot has only a finite past. Theorem 6 shows that we must take this into account in the theory; otherwise we might be tempted to simplify the theory by the fiction of an infinite past, as we did in Section 5.

6.2 INITIATION. Accordingly we shall hereafter assume (except when we indicate otherwise) that the past for our nerve nets goes back from p (the present) a finite time only, the first moment of which shall be 1 on our time scale. The range of the time variable in our logical formulas shall be the integers from 1 forward.

Now if in the net of Figure 10 $\mathcal{M}$ is quiet at time 1, we do have $M(p+1) \equiv (Et)_{t\leq p}N(t)$; in Figure 11 if $\mathcal{M}$ fires at time 1, $\overline{M(p+1)} \equiv (t)_{t\leq p}N(t)$; and in Figure 12 if $\mathcal{M}$ fires at time 1, $\overline{M(p+1)} \equiv (Et)_{t\leq p}N(t)$. Thus the nets of Figures 10 and 12 are able to remember that $\mathcal{N}$ has fired since their beginning by changing $\mathcal{M}$ from the state it had initially; while the net of Figure 11 is able to recognize that $\mathcal{N}$ has never failed to fire by preserving $\mathcal{M}$ in the state it had originally, as Householder and Landahl [1945] have commented [p. 109].

The nets in question only represent the events in question, when the inner neuron $\mathcal{M}$ has the state mentioned at time 1.

This again is the general rule for indefinite events, by Theorem 7 with Lemma 1. (Lemma 1 holds for finite past.)

We illustrate this now by showing that to represent the event $(t)_{t\leq p}N(t)$ at least one inner neuron must fire at time 1. For let the proposed representation be by a property of the state at time p (Lemma 1), i.e., solely of this state and not also of the value of p. Say $\mathcal{N}$ is the only input neuron. Were all the inner neurons quiet at $t = 1$, then with the input shown in the table of Figure 13 all inner neurons would be quiet at $t = 2$, so the state at $t = 2$ would be indistinguishable from that with Figure 14 at $t = 1$. Hence with Figure 13 the net would have the same state for $p = 2, 3, 4, \ldots$ as with Figure 14 for $p = 1, 2, 3, \ldots$, respectively, though with the former $(t)_{t\leq p}N(t)$ is false, with the latter true.

Accordingly in studying the representability of events, we shall hereafter not only choose a net for the purpose but also choose the state (firing or non-firing) of each inner neuron at time 1.

As the example of $(t)_{t\leq p}N(t)$ shows, for some events it will not be sufficient to have all the inner neurons quiet initially.

We are developing the theory of McCulloch-Pitts nerve nets as an illustrative case of the theory of finite automata. From the standpoint of the latter theory, one initial state is as reasonable as another. The alternative of excluding such events as $(t)_{t\leq p}N(p)$ from the class of representable events would be more awkward.

t	1	2	3	4	...
$\mathcal{N}$	0	1	1	1	...

FIGURE 13

t	1	2	3	4	...
$\mathcal{N}$	1	1	1	1	...

FIGURE 14

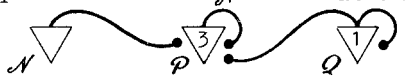
To one who feels that the firing of inner neurons at time 1 requires explanation under the McCulloch-Pitts laws of neural behavior, we need merely say that we have isolated certain input neurons $\mathcal{N}_1, \ldots, \mathcal{N}_k$ and a certain portion $t = 1, 2, 3, \ldots$ of time for the input for the events to be represented. We can go outside those neurons and that part of time to bring about any assumed state of our inner neurons at time $t = 1$. This is most simply accomplished by adding an extra input neuron $\mathcal{N}_{k+1}$, which is to fire at $t = 0$ and then only, and which is to have on each of the inner neurons we wish fired at $t = 1$ a number of excitatory endbulbs equal to the threshold of that neuron, but no other endbulbs.

A more complicated device, which requires an extra input neuron $\mathcal{K}$ but no extra moment of time, is illustrated in Figure 15 for the event $(t)_{t\leq p}N(p)$. The event is represented by firing $\mathcal{P}$ at $t = p + 2$, if $\mathcal{K}$ is fired, but all inner neurons are quiet, at $t = 1$. We can imagine $\mathcal{K}$ exposed to continual environmental stimulation which guarantees its firing at $t = 1$; but its firing at later times does not interfere with the representation. If $\mathcal{K}$ did not fire at $t = 1$, but first at some later time $t = u$, $\mathcal{P}$'s firing at $p + 2$ would represent that $\mathcal{N}$ had fired at all moments from u to p inclusive.[2]

6.3 DEFINITE EVENTS RECONSIDERED. An event is a partition of the class of all the possible inputs over the past (including the present) into two subclasses, those inputs for which the event occurs and those inputs for which

---

[2] McCulloch and Pitts consider the problem of "solving" nets with their initial state unspecified. To "solve" for a given inner neuron $\mathcal{P}$, say at time $p + 1$, means then to find for which inputs over time $1, \ldots, p$ and initial states of the inner neurons $\mathcal{P}$ will fire at time $p + 1$. In the following net, the necessary and sufficient condition that $\mathcal{P}$ fire at $p + 1$ is that $\mathcal{N}$ fire at all times $\leq p$ and both $\mathcal{P}$ and $\mathcal{Q}$ fire



$$P(p + 1) \equiv (t)_{t\leq p}N(t) \;\&\; P(1) \;\&\; Q(1).$$

at time 1. This seems to be a counterexample to the formula next after (9) on p. 126 of McCulloch-Pitts [1943], the proof of which we did not follow; for if we understand the formula correctly, it implies that the condition for firing of $\mathcal{P}$ should only require the existence of one neuron that fires initially. (Their 0 seems to be our 1.) This apparent counterexample discouraged us from further attempts to decipher Part III of McCulloch-Pitts [1943].

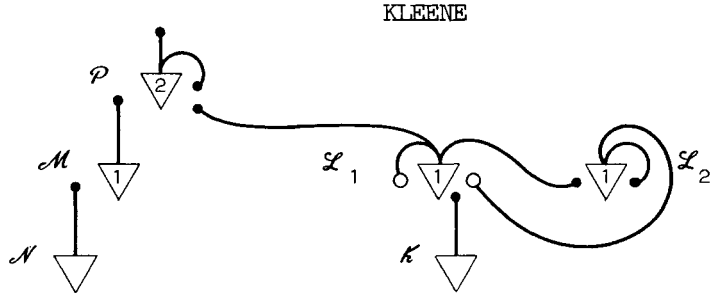


FIGURE 15

the event does not occur. The possible inputs on k neurons $\mathcal{N}_1, \ldots, \mathcal{N}_k$ are described by $k \times p$ tables of 0's and 1's with columns for $\mathcal{N}_1, \ldots, \mathcal{N}_k$ and rows for $t = p, \ldots, 1$. As p varies over all positive integers, these are all tables of 0's and 1's with k columns and any finite number of rows.

In Section 5 we used $k \times \ell$ tables to describe inputs over the last ℓ moments ending with the present.

Now that our time has an initial moment 1, we must be careful whenever we give a table with k columns and a finite number, say ℓ, of rows to describe an input, to make it clear whether we intend it to describe the input over the complete past (so $p = \ell$) or only over the last ℓ moments of the past (so $p \geq \ell$). In the one case we call the table initial, in the other non-initial. The table may be thought of as carrying a tag saying, respectively, $p = \ell$ or $p \geq \ell$. There was no necessity for this in Sections 4 and 5, as there tables referring to the complete past were infinite.

A definite event of length ℓ in Section 5 was one in which the partition of the inputs over the complete past is such that any two inputs which agree in the upper ℓ rows of their tables always fall into the same one of the subclasses. But now when $p < \ell$ there won't be ℓ rows in the table describing the input. For such a p, can the event occur? The convention we adopt is that the event shall not occur in this case. Thus the inputs of the first subclass for a definite event of length ℓ are those described by a set of non-initial $k \times \ell$ tables. If E_1 is the logical formula we used in Section 5 to describe a definite event, the event is now described by $E_1 \;\&\; p \geq \ell$. The negation of this is $\overline{E}_1 \vee p < \ell$, while the formula for the complementary definite event of length ℓ is $\overline{E}_1 \;\&\; p \geq \ell$, which is not equivalent, except for $\ell = 1$ when the "$\& p \geq \ell$" and "$\vee p < \ell$" are superfluous.

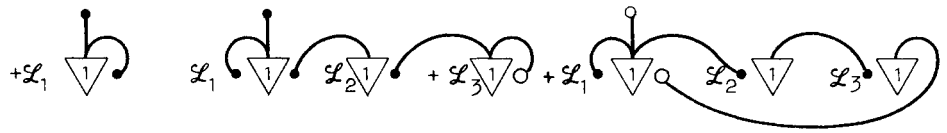
The identical event (written I_k or briefly I) which occurs no matter what the input (the second subclass of the partition being empty) is a definite event of length 1; the improper event (written $\overline{I}_k$ or $\overline{I}$) which never occurs (the first subclass being empty) can be considered as a definite event of length ℓ for every ℓ.

With the sole exception of the improper event, a given event can be definite of length ℓ for only one ℓ, and the set of the $k \times \ell$ tables (all of them non-initial) which describes it is unique. This was not the case in Section 5, as there a definite event of length ℓ was also definite of length m for each $m > \ell$; but now this would be absurd (except for the improper event) as the extra specification that $p \geq m$ would contradict that the event can occur for $p = \ell, \ldots, m - 1$.

The definite events we have just finished describing ($2^{2^{k\ell}}$ of them for a given k and ℓ) are those which arise most naturally from those considered in Section 5 by taking into account that now the past may not include ℓ moments.

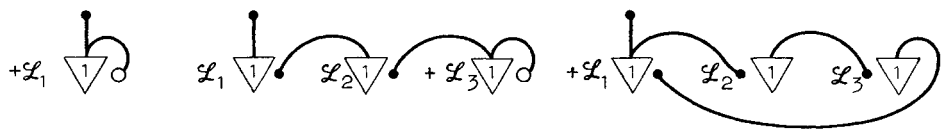
We now find it advantageous to introduce also a new kind of definite event on k neurons of length ℓ, by changing the specification for all the tables that $p \geq \ell$ to $p = \ell$; we do not include the improper event among these. These definite events we call <u>initial</u>. For a given k and ℓ, there are $2^{2^{k\ell}} - 1$ of them. An event can be an initial definite event for only one ℓ, and the set of the $k \times \ell$ tables (all of them initial) which describes it is unique. If E_1 & $p \geq \ell$ is a given non-initial definite event not improper, E_1 & $p = \ell$ is the corresponding initial definite event.

In Section 5 p entered the formulas for events only relatively, but now the events can refer to the value of p. This may seem somewhat unnatural; but, reversing the standpoint from which we were led to this in 6.1 and 6.2, if we are to analyze nerve nets in general, starting from arbitrary initial states, we are forced to give p an absolute status. This is illustrated in Figures 16-21, where the formula gives for each net the "solution" for L_1, i.e., the condition for its firing. The "+" indicates initial firing of the indicated neuron; inner neurons not bearing a "+" are initially quiet.



$L_1(p) \equiv p \geq 1.$ $L_1(p) \equiv p \geq 3.$ $L_1(p) \equiv p \leq 3.$

FIGURE 16 FIGURE 17 FIGURE 18



$L_1(p) \equiv p = 1.$ $L_1(p) \equiv p = 3.$ $L_1(p) \equiv [p \equiv 1 \pmod 3].$

FIGURE 19 FIGURE 20 FIGURE 21